# UNITED STATES PATENT AND TRADEMARK OFFICE

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/518,539 | 03/30/2005 | Hadi Harb | 04253 | 9563 |

23338    7590    09/22/2008
DENNISON, SCHULTZ & MACDONALD
1727 KING STREET
SUITE 105
ALEXANDRIA, VA 22314

| EXAMINER |
|---|
| LERNER, MARTIN |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2626 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 09/22/2008 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

1)☒ Responsive to communication(s) filed on <u>30 December 2004</u>.
2a)☐ This action is **FINAL**.      2b)☒ This action is non-final.
3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

4)☒ Claim(s) <u>1 to 33</u> is/are pending in the application.
    4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5)☐ Claim(s) _____ is/are allowed.
6)☒ Claim(s) <u>1 to 33</u> is/are rejected.
7)☐ Claim(s) _____ is/are objected to.
8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

## Application Papers

9)☒ The specification is objected to by the Examiner.
10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.
    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    a)☒ All   b)☐ Some * c)☐ None of:
    1.☒ Certified copies of the priority documents have been received.
    2.☐ Certified copies of the priority documents have been received in Application No. _____.
    3.☒ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
    * See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**
1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date <u>03/30/05</u>.
4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

# DETAILED ACTION

## *Specification*

1.      The abstract of the disclosure is objected to because it should be one paragraph only.  Correction is required.  See MPEP § 608.01(b).

2.      The disclosure is objected to because of the following informalities:

On page 6, lines 15 to 18, "hundredths of milliseconds" should be "hundreds of milliseconds".  Hundredths of milliseconds would correspond to microseconds, and would not be in the range of 0.5 to 2 seconds.

On page 6, lines 28 to 29, "noted after the DFT" should be "noted after as the DFT".

Appropriate correction is required.

3.      The disclosure is objected to because it contains embedded hyperlinks and/or other form of browser-executable code.  Applicants are required to delete the embedded hyperlinks and/or other form of browser-executable code.  See MPEP § 608.01.

Page 13, Line 26; Page 16, Line 2; and Page 17, Line 14, have embedded hyperlinks that need to be deleted.

## *Claim Rejections - 35 USC § 112*

4.      The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

5.      Claims 7, 13, 14, and 31 are rejected under 35 U.S.C. 112, second paragraph,

as being indefinite for failing to particularly point out and distinctly claim the subject

matter which applicant regards as the invention.

Claims 7, 13, and 31 contain the phrase "such as", which renders the claims

indefinite because it is unclear whether the limitations following the phrase are part of

the claimed invention.  See MPEP § 2173.05(d).

Claims 13  and 31 contain the phrase "for example", which renders the claims

indefinite because it is unclear whether the limitations following the phrase are part of

the claimed invention.  See MPEP § 2173.05(d).

Claim 14 is indefinite because it says that the standardization follows classifying

the sound signal.  However, classification of the sound signal is actually the last step, so

that it is incorrect to say that the standardization follows classifying the sound signal.


### Claim Rejections - 35 USC § 102

6.      The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public
> use or on sale in this country, more than one year prior to the date of application for patent in the United
> States.

7.      Claims 1 to 4, 6 to 9, 11 to 13, 18, 20 to 22, and 24 to 32 are rejected under 35

U.S.C. 102(b) as being anticipated by *Liu et al. ("Audio Feature Extraction and Analysis*

*for Scene Segmentation and Classification")*.

Regarding independent claims 1 and 24, *Liu et al.* discloses a method and

apparatus for classification of audio information, comprising:

"dividing the sound signal into temporal segments (T) having a specific duration"

– audio feature analysis involves dividing a sampled audio signal into over-lapping

frames of 512 samples (§2. Audio Feature Analysis: Pages 62 to 63: Figure 1);

"extracting the frequency parameters of the sound signal in each of the

temporal segments (T), by determining a series of values of the frequency spectrum at

a frequency range between a minimum frequency and a maximum frequency" –

frequency domain features are calculated over each audio frame by taking a short-time

Fourier transform, $S_i(\omega)$, for each $i$th frame; frequency ranges for four subbands range

from a minimum of 0 Hz to a maximum of 11025 Hz (§2. Audio Feature Analysis:

Frequency Domain Features: Page 66);

"assembling the parameters in time windows (F) having a specific duration

greater than the duration of the temporal segments" – sampled audio is divided into

clips of 1 second long, containing 22050 samples (§2. Audio Feature Analysis: Pages

62 to 63: Figure 1); clips having a duration of one second containing 22050 samples are

greater in duration than temporal segments of frames having a duration of 512 samples;

"extracting from each time window (F), characteristic components" – twelve clip-

level audio features are obtained, including nonsilence ratio, voice-or-music ratio, and

frequency bandwidth energy ratios (§2. Audio Feature Analysis: Page 68);

"and on the basis of the extracted characteristic components, and using a

classifier, identifying the sound class of the time windows (F) of the sound signal" –

semantic contain of short audio clips are characterized by a neural network classifier

from clip-level features (Abstract; §4. Audio-Based Scene Classification: Pages 71 to

74).

Regarding claim 2, *Liu et al.* discloses that clips of 1 second long contain 22050

samples, and frames contain 512 samples (§2. Audio Feature Analysis: Page 63: Figure

1); doing the math, (512 samples/22050 samples) seconds = 0.02 seconds = $20 \times 10^{-3}$

sec = 20 msec.

Regarding claims 3 to 4 and 25 to 26, *Liu et al.* discloses that frequency domain

features are obtained for each audio frame from a short-time Fourier transform (§2.

Audio Feature Analysis: Page 66); audio is sampled and digitized; thus, a Fourier

transform of a digital audio signal is equivalent to a "Discrete Fourier Transform"; a

Fourier transform is "an operation for transforming frequency parameters".

Regarding claims 6 and 27, *Liu et al.* discloses that clip-level features are

computed based on frame-level features for clips that are 1 second long (§2. Audio

Feature Analysis: Page 63: Figure 1).

Regarding claims 7 and 28, *Liu et al.* discloses extracting features from frames

including zero crossing rate ("silence crossing rate"), mean ("average"), and variance

(§2. Audio Feature Analysis: Page 64; §3.1 Mean and Variance Analysis: Page 68).

Regarding claim 8, *Liu et al.* discloses twelve audio features are obtained for

classification (§2. Audio Feature Analysis: Page 68).

Regarding claims 9 and 29, *Liu et al.* discloses, at least for frequency domain features, that subband energy ratios are used by dividing an energy in a subband by a total energy, where the subband ratios sum to 1 (§2. Audio Feature Analysis: Pages 66 to 67); thus, the energy in each subband is normalized relative to a total energy ("providing a standardization operation").

Regarding claims 11 and 30, *Liu et al.* discloses classification by both a nearest neighbor classifier and a neural network classifier (§4. Audio-Based Scene Classification: Page 72).

Regarding claim 12, *Liu et al.* discloses both a training phase and a testing phase from training data sets and testing data sets (§3. Feature Space Evaluation: Page 68; §4. Audio-Based Scene Classification: Page 71).

Regarding claims 13 and 31, *Liu et al.* discloses classification into voice or music by voice-or-music ratio (VMR) (§2. Audio Feature Analysis: Page 65); moreover, video is classified into characteristic moments of football, and uncharacteristic moments of advertisements (§4. Audio-Based Scene Classification: Pages 75 to 76: Figure 12(a)).

Regarding claims 18, 20, and 21, *Liu et al.* discloses classification into voice or music by voice-or-music ratio (VMR) ("identifying and monitoring the speech in a sound signal") ("identifying and monitoring music in a sound signal") ("determining if the sound signal contains speech or music") (§2. Audio Feature Analysis: Page 65).

Regarding claims 22 and 32, *Liu et al.* discloses scene segmentation provides labels for scenes of "football", "tv logo", "advertisement 1", "forecast", "news", *etc.* (§5. Scene Segmentation Using Audio Features: Pages 75 to 77: Figures 12(a) and 13(a)).

## Claim Rejections - 35 USC § 103

8.      The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

9.      Claims 5 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over

*Liu et al. ("Audio Feature Extraction and Analysis for Scene Segmentation and*

*Classification")* in view of *Burges et al.*

        Concerning claim 5, *Liu et al.* discloses that frequency bands may be divided

considering the perceptual property of the human ears into critical bands representing

cochlear filters in the human auditory model.  (§2. Audio Feature Analysis: Page 66)

Those skilled in the art know that a Mel scale and a Bark scale are the only ways to

represent critical bands in the human perceptual audio model.  Still, *Liu et al.* does not

expressly disclose a Mel scale.  However, *Burges et al.* teaches extracting features from

audio signals for use in classification, where it is stated that current audio classification,

segmentation and retrieval methods use heuristic features such as mel cepstra.

(Column 1, Lines 26 to 36)  It would have been obvious to one having ordinary skill in

the art to employ a Mel scale for extracting audio features as taught by *Burges et al.* in a

method and apparatus for audio feature extraction of *Liu et al.* because it is well known

that mel cepstra are employed to represent critical bands in a human auditory

perceptual model.

Concerning claim 19, *Liu et al.* discloses that audio clips are collected for five

scene classes including one or two male/female reporters. (§3. Feature Space

Evaluation) However, *Liu et al.* does not expressly say that male speech or female

speech is identified. Still, *Burges et al.* teaches that a corpus of audio examples

includes male and female talkers, which are examples of the kinds (classes) of audio to

be discriminated between. (Column 2, Lines 7 to 12)


10.     Claim 10 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Liu et al.*

*("Audio Feature Extraction and Analysis for Scene Segmentation and Classification")* in

view of *Huang et al.*

*Liu et al.* discloses features including a zero crossing rate, mean and variance,

and standardization of subband energy ratios. (§2. Audio Feature Analysis: Page 64,

and 66 to 68) However, *Liu et al.* does not expressly disclose standardization by

dividing components by a maximum value or dividing components by a constant fixed

after experimentation to obtain a value between 0.5 and 1. Still, *Huang et al.* teaches a

method and apparatus for segmenting a multi-media program based upon audio events,

where features include at least a volume dynamic range, which involves normalizing a

volume by a maximum volume in a clip. (Column 4, Lines 42 to 45) Moreover, it is

maintained that normalizing a feature so that it lies in a range of 0 to 1 is a common

expedient in audio processing, and normalization in a range between 0.5 and 1 is a

matter of "design choice", in an absence of unexpected results. *Huang et al.* suggests

advantages of a simpler process for identifying and indexing commercials in television

news programs. (Column 2, Lines 11 to 17) It would have been obvious to one having

ordinary skill in the art to normalize by dividing by a maximum value or by dividing by an

empirical value as suggested by *Huang et al.*, and by considerations of "design choice",

in a method and apparatus for audio feature extraction and analysis of *Liu et al.* for a

purpose of providing a simpler process for segmenting a multi-media program.

11.      Claims 14 to 17, 23, and 33 are rejected under 35 U.S.C. 103(a) as being

unpatentable over *Liu et al.* *("Audio Feature Extraction and Analysis for Scene*

*Segmentation and Classification")* in view of *Dagtas*.

Concerning claim 14, *Liu et al.* discloses classification into voice or music by

voice-or-music ratio (VMR) (§2. Audio Feature Analysis: Page 65), and extracting

features from frames including zero crossing rate ("silence crossing rate"), mean

("average"), and variance. (§2. Audio Feature Analysis: Page 64; §3.1 Mean and

Variance Analysis: Page 68) However, *Liu et al.* discloses a clip of 1 second long, but

not of two seconds long. (§2. Audio Feature Analysis: Page 63: Figure 1) Still, *Dagtas*

teaches a system and method for detecting highlights in a video program, where

segments equivalent to five seconds are used to compute average strengths of true

interesting events. (Column 7, Lines 27 to 31) An objective is to improve systems that

are capable of detecting interesting events in a video program. (Column 1, Lines 61 to

65) It would have been obvious to one having ordinary skill in the art to vary a length of

an audio clip from one second to two seconds as suggested by *Dagtas* in a method and

apparatus for audio feature extraction and analysis of *Liu et al.* for a purpose of

detecting interesting events in a video program.

Concerning claims 15 to 17, *Dagtas* teaches a system for detecting highlights

("strong moments") of a video program including a sports event ("a match") by

comparing an audio signal energy level to a threshold.  (Column 1, Lines 5 to 10;

Column 2, Lines 15 to 19)  A video playback device is capable of playing back only the

highlights ("a match summary") extracted from a video program (e.g., a sports program).

(Column 6, Lines 12 to 16)

Concerning claims 23 and 33, *Dagtas* teaches that an audio processor can

perform a textual search of audio track information ("searching for labels") to return a list

of candidates.  (Column 7, Lines 12 to 15)


### *Conclusion*

12.    The prior art made of record and not relied upon is considered pertinent to

Applicants' disclosure.

Foote, Whitman et al., Jiang et al., Gibbon et al., Trovato et al., and Dimitrova et

al. disclose related art.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to MARTIN LERNER whose telephone number is

(571)272-7608.  The examiner can normally be reached on 8:30 AM to 6:00 PM

Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.


/Martin  Lerner/
Primary Examiner
Art Unit 2626
September 18, 2008